

REMARKS

Claims 1-30 are all the claims presently pending in the application. Claims 1, 16, 23, 29 and 30 have been amended.

It is noted that the claim amendments herein or later are not made to distinguish the invention over the prior art or narrow the claims or for any statutory requirements of patentability. Further, Applicant specifically states that no amendment to any claim herein or later should be construed as a disclaimer of any interest in or right to an equivalent of any element or feature of the amended claim.

Applicant gratefully acknowledges that claims 12 and 24 would be allowable if rewritten in independent form. However, Applicant respectfully submits that all of the claims are allowable.

Claims 1-8, 10, 11, 13-23, 25 and 27-29 stand rejected under 35 U.S.C. §103(a) as being allegedly unpatentable over Ganguly et al. ("Proximal: A Database System for the Efficient Retrieval of Genetic Information", *Comput. Biol. Med.* Vol. 26, No 3, pp 199-207, 1996) (hereinafter "Ganguly") in view of Rigoutsos et al. ("Building Dictionaries of 1D and 3D Motifs by Mining The Unaligned 1D sequences of 17 Archaeal and Bacterial Genomes", 1999) (hereinafter "Rigoutsos '99"). **Claims 9, 26 and 30** stand rejected under 35 U.S.C. §103(a) as allegedly unpatentable over Ganguly and Rigoutsos '99, and further in view of Delcher et al. ("Improved Microbial Gene Identification with GLIMMER", 1999) (hereinafter "Delcher").

These rejections are respectfully traversed in the following discussion.

I. THE CLAIMED INVENTION

The claimed invention (e.g., as recited in claim 1) is directed to a system for identifying genes. The system includes a pattern database comprising patterns of amino acids, and an input device for inputting a genomic DNA sequence.

Importantly, the system further includes a processor which translates an open reading frame (ORF) of the DNA sequence into an amino acid translation, and locates in the amino acid translation occurrences of the patterns from the pattern database and determines whether the open

reading frame includes at least part of a putative gene based on the occurrences of the **patterns of amino acids located in the amino acid translation of the ORF**.

Conventional systems for identifying genes (e.g., putative genes) are either based on the use the statistics of DNA sequences, or the use of similarity searches to determine gene locations (Application at page 2, lines 7-22). However, these conventional methods have various problems which prevent them from efficiently identifying genes in a given DNA sequence (Application at page 3, line 19-page 4, line 21).

The claimed invention, on the other hand, includes a processor which determines whether the open reading frame includes at least part of a putative gene based on the occurrences of the **patterns of amino acids located in the amino acid translation of the ORF** (Application at Figure 3; page 12, line 10-page 13, line 14). The claimed invention may be considered as including the best characteristics of statistical approaches and database similarity searches, in identifying genes in a given DNA sequence (Application at page 6, lines 18-21).

II. THE ALLEGED PRIOR ART REFERENCES

A. Ganguly and Rigoutsos '99

The Examiner alleges that Ganguly would have been combined with Rigoutsos '99 to form the invention of claims 1-8, 10, 11, 13-23, 25 and 27-29. Applicant submits, however, that Ganguly would not have been combined with Rigoutsos '99 and even if combined, the combination would not teach or suggest each and every feature of the claimed invention.

Ganguly discloses a database system which allows complex sequence queries, interspersing similarity-type specifications with features such as constraints on the values of compositional parameters, defined features such as open reading frames and repeats and experimentally determined features such as DNA binding sites (Ganguly at page 200, lines 30-35).

Rigoutsos '99 shows that given large collections of amino acid sequences one can building dictionaries of amino acid sequence motifs from these collections and discusses a process for doing this. Attached hereto as Exhibit 1 is a flowchart illustrating the process of

Rigoutsos '99.

Applicant respectfully submits that contrary to the Examiner's allegations, Rigoutsos '99 does not teach or suggest a processor which "*determines whether said open reading frame includes at least part of a putative gene based on said occurrences of said **patterns of amino acids located in said amino acid translation of said ORF***", as recited, for example, in claim 1, and similarly recited in claims 16, 23 and 29.

As noted above, unlike conventional systems for identifying genes (e.g., putative genes) which are either based on the use of the statistics of DNA sequences, or the use of similarity searches to determine gene locations, the claimed invention includes a processor that determines whether an open reading frame includes at least part of a putative gene based on the occurrences of the **patterns of amino acids located in the amino acid translation of the ORF** (Application at Figure 3; page 12, line 10-page 13, line 14). The claimed invention may be considered as including the best characteristics of statistical approaches and database similarity searches, in identifying genes in a given DNA sequence (Application at page 6, lines 18-21).

Clearly, these features are not taught or suggested by the Ganguly or Rigoutsos '99. Indeed, the Examiner concedes on page 3 of the Office Action that Ganguly does not teach or suggest this feature. The Examiner alleges that Rigoutsos '99 discloses this feature, but the Examiner is incorrect.

That is, nowhere does Rigoutsos '99 teach or suggest determining whether an open reading frame includes at least part of a putative gene based on the occurrences of **patterns of amino acids located in the amino acid translation of the ORF**. In fact, Rigoutsos '99 does not even address or discuss the problem of finding genes anywhere in its presentation.

The Examiner attempts to rely on page 228, col. 2, lines 4-11 in Rigoutsos '99 to support his position. However, this is completely unreasonable.

In fact page 228, col. 2, lines 4-11 in Rigoutsos '99 merely reads (in pertinent part)
*"[o]ne of the seqlets that are discovered when we process the input database is
and is present in the following ten ORFs: Of these, gi_3328856 and
gi_3329230 are annotated as Fe-S oxidoreductases."*

That is, this short and simple passage does not teach or suggest locating occurrences of patterns from the pattern database in the amino acid translation of the ORF, and certainly does not teach or suggest determining whether an open reading frame includes at least part of a putative gene based on the occurrences of the **patterns of amino acids located in the amino acid translation of the ORF**.

Indeed, Applicant would again point out that the passage at page 228, col. 2, lines 4-11 in Rigoutsos '99 clearly does not teach or suggest locating occurrences of patterns from the pattern database in the amino acid translation. In fact, nowhere does this passage discuss "locating" anything, let alone locating occurrences of "patterns" in an amino acid translation of some DNA sequence. Indeed, in its entirety, the discussion of Rigoutsos '99 revolves around amino acid sequences: all of the discussed datasets comprise amino acid sequences. **Nowhere in Rigoutsos '99 does the discussion mention DNA sequences.**

Further, the Examiner seems to misunderstand the term "putative gene". Indeed, the Examiner seems to imply that this passage in Rigoutsos '99 teaches identifying a putative gene from the mere disclosure that two amino acid sequences (gi_3328856 and gi_3329230) are annotated as oxidoreductases.

This is clearly unreasonable.

Indeed, this passage simply states that the amino acid sequences with labels gi_3328856 and gi_3329230 respectively **were known at the time to be oxidoreductases**. Even if one focuses on gi_3328856 and gi_3329230, this very short passage does not teach or suggest how one can begin with the DNA sequence that corresponds to the DNA locations containing the DNA encodings for gi_3328856 (or for gi_3329230) and determine whether this DNA sequence indeed codes for a bona fide gene.

Thus, it is clear that Rigoutsos '99 merely discloses in this passage the knowledge that existed at the time in public databases that the two amino acid sequences, with identifiers gi_3328856 and gi_3329230, are **members of a certain class of enzymes**. Nowhere does Rigoutsos '99 describe how, where or why these two amino acids were so annotated, but simply reiterates knowledge that was available at the time about these two sequences.

Clearly, this passage does not teach or suggest a method for determining whether the DNA sequences that correspond to gi_3328856 and gi_3329230 respectively, represent the DNA coding of a bona fide gene. Additionally, nowhere in this passage or anywhere else, does Rigoutsos '99 teach or suggest locating occurrences of patterns from the pattern database in the amino acid translation, let alone determining whether an open reading frame (e.g., that has been presented as a DNA sequence) represents the encoding of at least part of a putative gene based on the identification of occurrences of amino acid **patterns within the amino acid sequence that represents the translation of the DNA sequence.**

Further, Applicant respectfully submits that the Examiner appears to be confused with a passage in Rigoutsos '99 at page 228, left col., first paragraph:

"[o]ne of the seqlets that are discovered when we process the input database is and is present in the following ten ORFs: Of these, gi_3328856 and gi_3329230 are annotated as Fe-S oxidoreductases."

Applicant submits that the examiner is likely misinterpreting this passage to mean that Rigoutsos '99 has taught how to annotate these two sequences and Rigoutsos '99 claims that these two sequences are Fe-S oxidoreductases. That is, the Examiner appears to be interpreting this phrase as "are annotated by Rigoutsos '99". This is clearly not correct.

In fact, nowhere does Rigoutsos '99 teach or suggest annotating seqlets as Fe-S oxidoreductases. Applicant respectfully submits that this simple passage could not possibly be considered to "teach" or "suggest" annotating these seqlets. Certainly, one of ordinary skill in the art could not read Rigoutsos '99 and **know how to annotate the seqlets as Fe-S oxidoreductases.** Indeed, nowhere does Rigoutsos '99 teach any method, process or criteria for annotating these seqlets.

In fact, this passage is simply indicating that **some third party has annotated these seqlets in the public databases as Fe-S oxidoreductases.** Indeed, the precise third party responsible for annotating these seqlets can be identified if one goes to the public records for these sequences and tracks down the bibliography related to these seqlets.

Thus, this passage should be interpreted "are annotated **by third parties** in the public databases as Fe-S oxidoreductases."

Moreover, even assuming (arguendo) that this passage teaches annotating the seqlets as Fe-S oxidoreductases, Applicant would point out that it would be the kind of "annotation" that claims a "functional behavior" and this kind of annotation has nothing whatsoever to do with identifying genes (e.g., "annotating a DNA sequence by claiming that it actually codes for a gene") which is a purpose of the claimed invention

Clearly, in the field of the claimed invention, the verb "annotate" is overloaded in terms of what it means. Thus, even if two people of ordinary skill in the art having a conversation agree to interpret "annotate" to mean "annotate a protein sequence by stating its putative or validated functional behavior", there is still ambiguity.

Moreover, Applicant would again point out that Rigoutsos '99 has nothing to do with "identifying genes". Indeed, this is clear from the title of Rigoutsos '99 which is "***Building Dictionaries Of 1D and 3D Motifs By Mining The Unaligned 1D Sequences of ... Genomes***". That is, Rigoutsos '99 deals with the process of building dictionaries of amino acid motifs using inputs comprising amino acid sequences. Rigoutsos '99 does not deal with, nor does it discuss, how one can **identify genes** by processing inputs comprising DNA sequences.

In fact, Rigoutsos '99 simply discloses a process that may be used to create a pattern database. Indeed, Rigoutsos '99 is summarized on page 224, col. 1 which states "[w]e report on the properties of the entries of these two dictionaries, the extent to which the 1D seqlets have instances in the Protein Data Bank, and the ramifications from the induced coverage. We also present several entries of the generated dictionaries and discuss uses". As is clear from this summary, Rigoutsos '99 has nothing to do with determining whether an open reading frame includes at least part of a putative gene based on the occurrences of the **patterns of amino acids located in the amino acid translation of the ORF**.

In fact, as clearly indicated in the flowchart illustrated in the attached Exhibit 1, Rigoutsos '99 has nothing to do with identifying genes and clearly does not teach or suggest the claimed invention. Indeed, **nowhere in Rigoutsos '99 is the term "DNA" or "DNA sequence"**

even used. Instead, Rigoutsos '99 only refers to and uses inputs including amino acid sequences. Thus, it is clear that contrary to the Examiner's allegations Rigoutsos '99 does not teach or suggest a method for inputting the DNA sequence that corresponds to an open reading frame and determining whether the DNA sequence represents a bona fide gene by examining whether the amino acid translation of this DNA sequence contains amino acid patterns from the pattern database.

Therefore, Applicant submits that these alleged references would not have been combined and even if combined, the combination would not teach or suggest each and every element of the claimed invention. Therefore, the Examiner is respectfully requested to withdraw this rejection.

B. Delcher

The Examiner alleges that Ganguly and Rigoutsos would have been further combined with Delcher to form the invention of claims 9, 26 and 30. Applicant submits, however, that these alleged references would not have been combined and even if combined, the combination would not teach or suggest each and every feature of the claimed invention.

Applicant respectfully submits that these references would not have been combined as alleged by the Examiner. Indeed, these references are unrelated, and no person of ordinary skill in the art would have considered combining these disparate references, absent impermissible hindsight.

In fact, these references clearly do not teach or suggest their combination. Therefore, Applicant respectfully submits that one of ordinary skill in the art would not have been so motivated to combine the references as alleged by the Examiner. Therefore, the Examiner has failed to make a prima facie case of obviousness.

Moreover, contrary to the Examiner's allegations, neither Ganguly, nor Rigoutsos nor Delcher, nor any combination thereof teaches or suggests a processor that "*determines whether said open reading frame includes at least part of a putative gene based on said occurrences of said patterns of amino acids located in said amino acid translation of said ORF*", as recited, for example, in claim 1, and similarly recited in claims 16, 23 and 29.

Clearly, Delcher does not teach or suggest these features of the claimed invention.

Indeed, Applicant would again point out that Delcher is unrelated to the claimed invention. In fact, Delcher uses nucleotide sequences for training, uses sequences from archaeal and bacterial genomes only for training, builds a Hidden Markov Model (HMM) from carefully-selected sequences, and uses the HMM to query nucleic acid sequences and compare the expectation in nucleotide sequence space as captured by the HMM with what is encountered in the input at hand;

Further, in Delcher, a comparison is done in the space of nucleotide sequences, and if its nucleotide sequences matches the expectation in terms of nucleotide sequence composition as the latter is captured by the HMM, then an ORF is reported as a putative gene.

That is, nowhere does Delcher teach or suggest determining whether an open reading frame includes at least part of a putative gene based on the occurrences of the **patterns of amino acids located in the amino acid translation of the ORF. Therefore, Delcher does not make up for the deficiencies in Ganguly and Rigoutsos '99.**

Therefore, Applicant submits that these alleged references would not have been combined and even if combined, the combination would not teach or suggest each and every element of the claimed invention. Therefore, the Examiner is respectfully requested to withdraw this rejection.

III. FORMAL MATTERS AND CONCLUSION

In view of the foregoing, Applicant submits that claims 1-30, all the claims presently pending in the application, are patentably distinct over the prior art of record and are in condition for allowance. The Examiner is respectfully requested to pass the above application to issue at the earliest possible time.

Should the Examiner find the application to be other than in condition for allowance, the Examiner is requested to contact the undersigned at the local telephone number listed below to discuss any other changes deemed necessary in a telephonic or personal interview.

Serial No. 10/059,421
Docket No. YOR920010126US2

17

The Commissioner is hereby authorized to charge any deficiency in fees or to credit any overpayment in fees to Assignee's Deposit Account No. 50-0510.

Respectfully Submitted,

Date: March 6, 2007

/Phillip E. Miller/
Phillip E. Miller, Esq.
Registration No. 46,060

McGinn IP Law Group, PLLC
8321 Old Courthouse Road, Suite 200
Vienna, VA 22182-3817
(703) 761-4100
Customer No. 21254